WE CLAIM:

1.   An apparatus for determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

   means for receiving a first sequence of sub-word units representative of a first one of said at least two words and for receiving a second sequence of sub-word units representative of a second one of said at least two words;

   means for aligning and for comparing sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units; and

   means for determining a sequence of sub-word units representative of the received words in dependence upon the aligned pairs of sub-word units determined by said aligning and comparing means.

2.   An apparatus according to claim 1, wherein said determining means is operable to determine said sequence of sub-word units by determining, for each aligned pair of sub-word units, a sub-word unit that it confusingly

similar to the first and second sub-word units of the aligned pair.

3. An apparatus according to claim 2, wherein said determining means comprises:

first comparing means for comparing, for each aligned pair, the first sequence sub-word unit in the aligned pair with each of a plurality of sub-word units taken from a set of predetermined sub-word units, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit and the respective sub-word units of the set;

second comparing means for comparing, for each aligned pair, the second sequence sub-word unit in the aligned pair with each of said plurality of sub-word units from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit and the respective sub-word units of the set;

means for combining the comparison scores obtained when comparing the first and second sequence sub-word units in the aligned pair with the same sub-word unit from the set, to generate a plurality of combined comparison scores;

third comparing means for comparing, for each aligned pair, the combined comparison scores generated by said combining means for the aligned pair; and

means for determining, for each aligned pair of sub-word units, a sub-word unit representative of the sub-word units in the aligned pair in dependence upon a comparison result output by said third comparing means for the aligned pair.

4.   An apparatus according to claim 3, wherein said first and second comparing means are operable to compare the first sequence sub-word unit and the second sequence sub-word unit respectively with each of the sub-word units in said set of sub-word units.

5.   An apparatus according to claim 3, wherein said first and second comparing means are operable to provide comparison scores which are indicative of a probability of confusing the corresponding sub-word unit taken from the set of predetermined sub-word units as the sub-word unit in the aligned pair.

6.   An apparatus according to claim 5, wherein said combining means is operable to combine the comparison scores in order to multiply the probabilities of

confusing the corresponding sub-word unit taken from the
set as the sub-word units in the aligned pair.

7.   An apparatus according to claim 6, wherein each of
said sub-word units in said set of predetermined sub-word
units has a predetermined probability of occurring within
a sequence of sub-word units and wherein said combining
means is operable to weigh each of said combined
comparison scores in dependence upon the respective
probability of occurrence for the sub-word unit of the
set used to generate the combined comparison score.

8.   An apparatus according to claim 7, wherein said
combining means is operable to combine said comparison
scores by calculating:

$$P\left(d_i^1\middle|p_r\right) P\left(d_j^2\middle|p_r\right) P\left(p_r\right)$$

where d_i^1 and d_j^2 are an aligned pair of first and
second sequence sub-word units respectively; $P(d_i|p_r)$ is
the comparison score output by said first comparing means
and is representative of the probability of confusing set
sub-word unit p_r as first sequence sub-word unit d_i^1;
$P(d_j^2|p_r)$ is the comparison score output by said second
comparing means and is representative of the probability

of confusing set sub-word unit p_r as second sequence sub-word unit d^2_j; and $P(p_r)$ is a weight which represents the probability of set sub-word unit p_r occurring in a sequence of sub-word units.

5

9. An apparatus according to claim 8, wherein said third comparing means is operable to identify the set sub-word unit p_r which gives the maximum combined comparison score and wherein said determining means is

10 operable to determine said sub-word unit representative of the sub-word units in the aligned pair as being the sub-word unit which provides the maximum combined comparison score.

15 10. An apparatus according to claim 6, wherein said comparison scores represent log probabilities and wherein said combining means is operable to multiply said probabilities by adding the respective comparison scores.

20 11. An apparatus according to claim 3, wherein each of the sub-word units in said first and second sequences of sub-word units belong to said set of predetermined sub-word units and wherein said first and second comparing means are operable to provide said comparison scores

using predetermined data which relate the sub-word units in said set to each other.

12. An apparatus according to claim 11, wherein said predetermined data comprises, for each sub-word unit in the set of sub-word units, a probability for confusing that sub-word unit with each of the other sub-word units in the set of sub-word units.

13. An apparatus according to claim 1, wherein said aligning and comparing means comprises dynamic programming means for aligning said first and second sequences of sub-word units using a dynamic programming technique.

14. An apparatus according to claim 13, wherein said dynamic programming means is operable to determine an optimum alignment between said first and second sequences of sub-word units.

15. An apparatus according to claim 1, wherein each of said sub-word units represents a phoneme.

16. An apparatus according to claim 1, wherein said receiving means is operable to receive a third sequence

of sub-word unit representative of a third one of the words output by said word recognition unit and wherein said aligning and comparing means is operable to simultaneously align and compare the sub-word units of the first, second and third sequences of sub-word units.

17. An apparatus according to claim 1, wherein said receiving means is operable to receive a third sequence of sub-word units representative of a third one of said words output by said recognition unit and wherein aligning and comparing means is operable to align and compare two sequences of sub-word units at a time.

18. An apparatus according to claim 1, further comprising a word to sub-word unit dictionary which is operable to receive the words output by said word recognition unit and to generate therefrom said sequences of sub-word units.

19. An apparatus according to claim 1, further comprising means for annotating a data file using the sub-word units output by said determining means.

20. An apparatus according to claim 19, wherein said annotating means is operable to annotate said data file

using said sub-word units and said words output by said word recognition unit.

21.   An apparatus according to claim 20, wherein said sequence of sub-word units and said words are combined to form annotation data for the data file.

22.   An apparatus according to claim 19, wherein said data file is one of: an audio data file, a video data file, an image data file or a text data file.

23.   An apparatus according to claim 1, wherein said word recognition unit comprises a speech recognition system.

24.   An apparatus according to claim 1, wherein said word recognition unit comprises a handwriting recognition system.

25.   An apparatus for determining a sequence of sub-word units representative of at least two words, the apparatus comprising:

      means for receiving a first sequence of sub-word units representative of a first word and for receiving a second sequence of sub-word units representative of a second word;

means for aligning sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units; and

means for determining a sequence of sub-word units representative of the first and second sequences of sub-word units by determining, for each aligned pair of sub-word units, a sub-word unit that is confusingly similar to the first and second sub-word units of the aligned pair.

26. An apparatus for determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving the words output by the word recognition unit;

means for generating a sequence of sub-word units representative of each of the received words;

means for aligning and comparing the sub-word units from each generated sequence of sub-word units to identify a number aligned groups of sub-word units; and

means for determining a sequence of sub-word units representative of the received words in dependence upon the aligned groups of sub-word units determined by said aligning and comparing means.

27. An apparatus for determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

5   means for receiving the words output by the word recognition unit;

means for generating a sequence of sub-word units representative of each of the received words;

means for aligning and comparing the sub-word units from each generated sequence of sub-word units to identify a number aligned groups of sub-word units; and

means for determining a sequence of sub-word units representative of the first and second sequences of sub-word units by determining, for each of the sub-word units in an aligned group, a sub-word unit that it confusingly similar to the sub-word units of the group.

28. An apparatus for determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving a first sequence of sub-word units representative of a first one of said at least two words and for receiving a second sequence of sub-word

units representative of a second one of said at least two words;

means for aligning and for comparing sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units;

first comparing means for comparing, for each aligned pair, the first sequence sub-word unit in the aligned pair with each of a plurality of sub-word units taken from a set of predetermined sub-word units, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit and the respective sub-word units of the set;

second comparing means for comparing, for each aligned pair, the second sequence sub-word unit in the aligned pair with each of said plurality of sub-word units from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit and the respective sub-word units of the set;

means for combining the comparison scores obtained when comparing the first and second sequence sub-word units in the aligned pair with the same sub-word unit

from the set, to generate a plurality of combined comparison scores;

third comparing means for comparing, for each aligned pair, the combined comparison scores generated by said combining means for the aligned pair; and

means for determining, for each aligned pair of sub-word units, a sub-word unit representative of the sub-word units in the aligned pair in dependence upon a comparison result output by said third comparing means for the aligned pair.

29. A method of determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving a first sequence of sub-word units representative of a first one of said at least two words;

receiving a second sequence of sub-word units representative of a second one of said at least two words;

aligning and comparing sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units; and

determining a sequence of sub-word units representative of the received sequences of sub-word

units in dependence upon the aligned pairs of sub-word units determined in said aligning and comparing step.

30. A method according to claim 29, wherein said determining step determines said sequence of sub-word units by determining, for each aligned pair of sub-word units, a sub-word unit that is confusingly similar to the first and second sub-word units of the aligned pair.

31. A method according to claim 29, wherein said determining step comprises:

    a first comparing step of comparing, for each aligned pair, the first sequence sub-word unit in the aligned pair with each of a plurality of sub-word units taken from a set of predetermined sub-word units, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit and the respective sub-word units of the set;

    a second comparing step of comparing, for each aligned pair, the second sequence sub-word unit in the aligned pair with each of said plurality of sub-word units from the set, to provide a further corresponding plurality of comparison scores representative of the

similarities between said second sequence sub-word unit and the respective sub-word units of the set;

combining the comparison scores obtained when comparing the first and second sequence sub-word units in the aligned pair with the same sub-word unit from the set, to generate a plurality of combined comparison scores;

a third comparing step of comparing, for each aligned pair, the combined comparison scores generated in said combining step for the aligned pair; and

determining, for each aligned pair, a sub-word unit representative of the sub-word units in the aligned pair in dependence upon a comparison result output from said third comparing step for the aligned pair.

32. A method according to claim 31, wherein said first and second comparing steps compare the first sequence sub-word unit and the second sequence sub-word unit respectively with each of the sub-word units in said set of sub-word units.

33. A method according to claim 31, wherein said first and second comparing steps provide comparison scores which are indicative of a probability of confusing the corresponding sub-word unit taken from the set of

predetermined sub-word units as the sub-word unit in the
aligned pair.

34. A method according to claim 33, wherein said
combining step combines the comparison scores in order to
multiply the probabilities of confusing the corresponding
sub-word unit taken from the set as the sub-word units in
the aligned pair.

35. A method according to claim 34, wherein each of said
sub-word units in said set of predetermined sub-word
units has a predetermined probability of occurring within
a sequence of sub-word units and wherein said combining
step weighs each of said combined comparison scores in
dependence upon the respective probability of occurrence
for the sub-word unit of the set used to generate the
combined comparison score.

36. A method according to claim 35, wherein said
combining step combines said comparison scores by
calculating:

$$P\left(d_i^1|p_r\right) P\left(d_j^2|p_r\right) P\left(p_r\right)$$

where d^1_i and d^2_j are an aligned pair of first and second sequence sub-word units respectively; $P(d_i|p_r)$ is the comparison score output by said first comparing means and is representative of the probability of confusing set sub-word unit p_r as first sequence sub-word unit d^1_i; $P(d^2_j|p_r)$ is the comparison score output by said second comparing means and is representative of the probability of confusing set sub-word unit p_r as second sequence sub-word unit d^2_j; and $P(p_r)$ is a weight which represents the probability of set sub-word unit p_r occurring in a sequence of sub-word units.

37.   A method according to claim 36, wherein said third comparing step identifies the set sub-word unit p_r which gives the maximum combined comparison score and wherein said determining step determines said sub-word unit representative of the sub-word units in the aligned pair as being the sub-word unit which provides the maximum combined comparison score.

38.   A method according to claim 34, wherein said comparison scores represent log probabilities and wherein said combining step multiplies said probabilities by adding the respective comparison scores.

39. A method according to claim 31, wherein each of the sub-word units in said first and second sequences of sub-word units belong to said set of predetermined sub-word units and wherein said first and second comparing steps provide said comparison scores using predetermined data which relate the sub-word units in said set to each other.

40. A method according to claim 39, wherein said predetermined data comprises, for each sub-word unit in the set of sub-word units, a probability for confusing that sub-word unit with each of the other sub-word units in the set of sub-word units.

41. A method according to claim 29, wherein said aligning and comparing step uses a dynamic programming technique to align said first and second sequences of sub-word units.

42. A method according to claim 41, wherein said dynamic programming technique determines an optimum alignment between said first and second sequences of sub-word units.

43. A method according to claim 29, wherein each of said sub-word units represents a phoneme.

44. A method according to claim 29, further comprising receiving a third sequence of sub-word unit representative of a third one of the words output by said word recognition unit and wherein said aligning and comparing step simultaneously aligns and compares the sub-word units of the first, second and third sequences of sub-word units.

45. A method according to claim 29, further comprising receiving a third sequence of sub-word units representative of a third one of said words output by said recognition unit and wherein said aligning and comparing step aligns and compares two sequences of sub-word units at a time.

46. A method according to claim 29, further comprising the step of using a word to sub-word unit dictionary to convert the words output by said word recognition unit into said sequences of sub-word units.

47. A method according to claim 29, further comprising the step of annotating a data file using the sub-word units output by said determining step.

48. A method according to claim 47, wherein said annotating step annotates said data file using said sub-word units and said words output by said word recognition unit.

49. A method according to claim 48, wherein said sequence of sub-word units and said words are combined to form annotation data for the data file.

50. A method according to claim 47, wherein said data file is one of: an audio data file, a video data file, an image data file, or a text data file.

51. A method according to claim 29, wherein said word recognition unit comprises a speech recognition system.

52. A method apparatus according to claim 29, wherein said word recognition unit comprises a handwriting recognition system.

53.   A method of determining a sequence of sub-word units
representative   of   at   least   two   words,   the   method
comprising the steps of:

receiving   a   first   sequence   of   sub-word   units
representative of a first word;

receiving   a   second   sequence   of   sub-word   units
representative of a second word;

aligning sub-word units of the first sequence with
sub-word units of the second sequence to form a number of
aligned pairs of sub-word units; and

determining   a   sequence   of   sub-word   units
representative of the first and second sequences of sub-
word units by determining, for each aligned pair of sub-
word units, a sub-word unit that is confusingly similar
to the first and second sub-word units of the aligned
pair.

54.   A method of determining a sequence of sub-word units
representative of at least two words output by a word
recognition unit in response to a common input word to be
recognised, the method comprising the steps of:

receiving the words output by the word recognition
unit;

generating   a   sequence   of   sub-word   units
representative of each of the received words;

aligning and comparing the sub-word units from each generated sequence of sub-word units to identify a number of aligned groups of sub-word units; and

determining a sequence of sub-word units representative of the received words in dependence upon the aligned groups of sub-word units determined in said aligning and comparing step.

55. A method of determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the method comprising the steps of:

receiving the words output by the word recognition unit;

generating a sequence of sub-word units representative of each of the received words;

aligning and comparing the sub-word units from each generated sequence of sub-word units to identify a number of aligned groups of sub-word units; and

determining a sequence of sub-word units representative of the first and second sequences of sub-word units by determining, for each of the sub-word units in an aligned group, a sub-word unit that is confusingly similar to the sub-word units of the group.

56. A method of determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving a first sequence of sub-word units representative of a first one of said at least two words;

receiving a second sequence of sub-word units representative of a second one of said at least two words;

aligning and comparing sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units;

a first comparing step of comparing, for each aligned pair, the first sequence sub-word unit in the aligned pair with each of a plurality of sub-word units taken from a set of predetermined sub-word units, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit and the respective sub-word units of the set;

a second comparing step of comparing, for each aligned pair, the second sequence sub-word unit in the aligned pair with each of said plurality of sub-word units from the set, to provide a further corresponding plurality of comparison scores representative of the

similarities between said second sequence sub-word unit and the respective sub-word units of the set;

combining the comparison scores obtained when comparing the first and second sequence sub-word units in the aligned pair with the same sub-word unit from the set, to generate a plurality of combined comparison scores;

a third comparing step of comparing, for each aligned pair, the combined comparison scores generated in said combining step for the aligned pair; and

determining, for each aligned pair, a sub-word unit representative of the sub-word units in the aligned pair in dependence upon a comparison result output from said third comparing step for the aligned pair.

57. A computer readable medium storing computer executable process steps to perform a method of determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the process steps comprising the steps of:

receiving a first sequence of sub-word units representative of a first one of said at least two words;

receiving a second sequence of sub-word units representative of a second one of said at least two words;

aligning and comparing sub-word units of the first sequence with sub-word units of the second sequence to form a number of aligned pairs of sub-word units; and

determining a sequence of sub-word units representative of the received sequences of sub-word units in dependence upon the aligned pairs of sub-word units determined in said aligning and comparing step.

58. A computer executable program for controlling a processor to perform a method of determining a sequence of sub-word units representative of at least two words output by a word recognition unit in response to a common input word to be recognised, the program comprising:

code for receiving a first sequence of sub-word units representative of a first one of said at least two words;

code for receiving a second sequence of sub-word units representative of a second one of said at least two words;

code for aligning and comparing sub-word units of the first sequence with sub-word units of the second

sequence to form a number of aligned pairs of sub-word
units; and

    code for determining a sequence of sub-word units
representative of the received sequences of sub-word
5   units in dependence upon the aligned pairs of sub-word
units determined in said aligning and comparing step.